H–Bases I: The Foundation

H. Michael Möller and Thomas Sauer

Abstract. The H–basis concept allows an investigation of multivariate polynomial spaces degree by degree. In this paper, we mention its connection to the Gröbner basis concept, characterize H–bases, show how to construct them, and present a procedure for simplifying polynomials to their normal forms. Applications will be given in [8].

§1. Introduction

We consider Π, the ring of polynomials in $x_1, \ldots, x_n$ with coefficients from an infinite field $\mathbb{K}$, i.e. $\Pi = \mathbb{K}[x_1, \ldots, x_n]$, and the subsets Π_d of all polynomials of degree at most d. In many applications, one is interested in getting a basis or a generating set for the linear vector space $I \cap \Pi_d$, where $I \subseteq \Pi$ is an ideal. Having an H–basis $\{f_1, \ldots, f_s\}$ for I, then the set of all $p_i \cdot f_i$ with $p_i \in \Pi_{d-deg(f_i)}$, $i = 1, \ldots, s$, generates $I \cap \Pi_d$ as a linear vector space. Thus the H–basis concept is a tool for transforming a non-linear problem in Π into a problem in one (or in a series of) finite dimensional linear space(s) Π_d.

H–bases were introduced first by Macaulay [4]. His original motivation was the transformation of systems of polynomial equations into simpler ones. The power of this concept was not really understood, presumably because of the lack of facilities for symbolic computations. When Computer Algebra Systems came up, Gröbner bases (G–bases for short) were used instead of H–bases. These bases, originally invented by Buchberger [2] for computing multiplication tables for factor rings, are now also applied for simplifying some problems in Numerical Analysis, see [5].

The G–bases give generating systems not to $I \cap \Pi_d$ but to $I \cap \mathcal{F}_i$, where $\mathcal{F}_i \subset \Pi$ is a linear vector space of dimension i, and $\mathcal{F}_i \subset \mathcal{F}_{i+1}$ for all i and $\Pi = \cup_{i \geq 0} \mathcal{F}_i$. This finer decomposition has some drawbacks. For instance if an ideal is invariant under an affine symmetry group, its G–bases are typically not invariant. Since the spaces Π_d are invariant under affine symmetry groups, H–bases do not destroy such symmetries.

Many of the problems in applications which can be solved by Gröbner techniques can also be treated successfully with H–bases. In [7] we gave an

Curve and Surface Fitting: Saint-Malo 1999
Albert Cohen, Christophe Rabut, and Larry L. Schumaker (eds.), pp. 325–332.

ISBN 0-8265-1357-3.

overview of such problems. In the present paper, we describe briefly the underlying concept of grading rings, which leads to G- and H-bases, and present some properties characterizing H-bases. In contrast to [7], where we only gave a class of examples of H-bases, we present here the construction of an H-basis for zero-dimensional ideals I. A useful tool for our procedure is the so called normal form mapping NF, presented in Section 4, which projects Π orthogonally to the ideal I provided an H-basis of I is given. In [8] we show how these normal forms can be applied in numerical applications.

§2. H-bases and G-bases

In ring theory, rings can be graded by an ordered monoid, i.e. by an abelian semigroup Γ with addition $+$ and total ordering $\prec$ satisfying

$$\gamma_1 \prec \gamma_2 \;\Rightarrow\; \gamma_0 + \gamma_1 \prec \gamma_0 + \gamma_2, \quad \forall \gamma_0, \gamma_1, \gamma_2 \in \Gamma.$$

There are two major examples for grading Π by an ordered monoid Γ:

$$\Pi = \bigoplus_{\gamma \in \Gamma} \Pi_\gamma^{(\Gamma)}, \quad \Pi_{\gamma_1}^{(\Gamma)} \Pi_{\gamma_2}^{(\Gamma)} \subseteq \Pi_{\gamma_1+\gamma_2}^{(\Gamma)} \; \forall \gamma_1, \gamma_2 \in \Gamma.$$

The first one is the H-grading with $\Gamma := \mathbb{N}_o$,

$$\Pi_\gamma^{(\Gamma)} := \{p \in \Pi \mid p \textit{ homogeneous of order } \gamma\}.$$

The ordering of $\Gamma = \mathbb{N}_o$ is the natural one. The second example for gradings is the G-grading, where $\Gamma := \mathbb{N}_o^n$ and

$$\Pi_{(\gamma_1,\ldots,\gamma_n)}^{(\Gamma)} := \{c x_1^{\gamma_1} \cdots x_n^{\gamma_n} \mid c \in \mathbb{K}\}.$$

$\Gamma = \mathbb{N}_o^n$ is ordered by an admissible term ordering,

$$0 \preceq i,\; i \prec j \Rightarrow i + k \prec j + k.$$

Since the decomposition of Π into the sets $\Pi_\gamma^{(\Gamma)}$ is a direct sum, every $f \in \Pi$ has a unique representation $f = \sum f_\gamma$. The maximal γ with $f_\gamma \neq 0$ is called the maximal part of $f \neq 0$, $M^{(\Gamma)}(f)$ for short. It is also called the maximal form in the H-case, or leading monomial in the G-case. In the G-case, $M^{(\Gamma)}(f) = lc(f)lt(f)$, where $lc(f) \in \mathbb{K}$ is the leading coefficient and $lt(f) = x_1^{i_1} \cdots x_n^{i_n}$ the leading term. The maximal form of $f \neq 0$ is also denoted by $M_H(f)$.

Definition 1. $\{p_1, \ldots, p_m\} \subset I$ *is called a* basis *of an ideal* $I \subseteq \Pi$, *briefly* $I = \langle p_1, \ldots, p_m \rangle$, *if* $\forall\, p \in I$

$$\exists g_1, \ldots, g_m \in \Pi \;:\; p = \sum_{k=1}^{m} g_k p_k.$$

It is also a G-basis *or* H-basis *resp. if* $g_1, \ldots, g_m$ *satisfy in addition*

$$\max_{k=1}^{m} lt(g_k)lt(p_k) = lt(p) \quad (G - basis),$$

$$or \max_{k=1}^{m} \deg(g_k p_k) = \deg(p) \quad (H - basis).$$

Theorem 1. *Let* $I = \langle p_1, \ldots, p_m \rangle$. *Then* $\{p_1, \ldots, p_m\}$ *is an H–basis (G–basis resp.) if and only if the least ideal containing all* $M_H(f)$, $0 \neq f \in I$ *(or all* $lt(f)$, $0 \neq f \in I$ *resp.) is generated by* $M_H(p_1), \ldots, M_H(p_m)$ *(or by* $lt(p_1), \ldots, lt(p_m)$ *resp.).*

This theorem, which holds *mutatis mutandis* for arbitrary graded rings, is proved for instance in [6]. An immediate consequence of it is that every ideal $I \neq (0)$ has an H– and a G–basis.

G–bases are now a standard tool in Computer Algebra. They are covered by nearly all textbooks, and are contained in almost all Computer Algebra Systems. The grading by one-dimensional linear spaces $\Pi_\gamma^{(\Gamma)}$ often reduces the computation to solving a series of one-dimensional problems. On the other hand, the construction of G–bases is often difficult or even impossible because of the high complexity of Buchberger's algorithm for computing G–bases. In addition, in many applications the G–bases allows only little insight into the structure of a solution by the artificial ordering term by term.

§3. Characterization of H–bases and Normal Forms

Macaulay introduced H–bases using homogenizations and dehomogenizations of polynomials. The name *H–basis* originates from the first letter of homogenization.

Definition 2. *Let* $f \in \mathbb{K}[x_1, \ldots, x_n]$ *have degree* d,

$$f = \sum_{i=0}^{d} f_i, \quad f_i \text{ homogeneous of degree } i, \quad f_d \neq 0.$$

Then introducing a new variable x_0, *the* homogenization *of* f *is a homogeneous degree* d *polynomial in* $\mathbb{K}[x_0, x_1, \ldots, x_n]$,

$$\Phi(f) := \sum_{i=0}^{d} x_0^{d-i} f_i.$$

A homogeneous $F \in \mathbb{K}[x_0, x_1, \ldots, x_n]$ *can be dehomogenized to an* $f \in \Pi$ *by* $x_0 = 1$.

For more details on homogenizations and their connection to projective coordinates, we refer to [3].

Theorem 2. (Macaulay [4]). *Let $I = \langle h_1, \ldots, h_s \rangle$. Then the following statements are equivalent*

1) *The least ideal containing all $\Phi(h)$, $0 \neq h \in I$, is $\langle \Phi(h_1), \ldots, \Phi(h_s) \rangle$.*
2) $x_0 F \in \langle \Phi(h_1), \ldots, \Phi(h_s) \rangle \Rightarrow F \in \langle \Phi(h_1), \ldots, \Phi(h_s) \rangle$.
3) *$\{h_1, \ldots, h_s\}$ is an H–basis of I.*

The power of the G–basis concept is mainly based on the possibility of reducing a polynomial to a simpler one by subtracting suitable multiples of elements of the G–basis. A consequent application of this reduction strategy gives the so called normal form, in a sense the simplest polynomial obtainable by the reductions. We translated this technique to H–bases in [7], and give here for consistency a short résumé of the main results.

Definition 3. *We denote by $\Pi_d^{(H)}$ the space of all homogeneous degree d polynomials for $d \in \mathbb{N}_o$ and $\Pi_d^{(H)} := \{0\}$ for $d < 0$. Let $h_1, \ldots, h_s \in \Pi$. Then we define a finite dimensional linear subspace of $\Pi_d^{(H)}$ by*

$$V_d(h_1, \ldots, h_s) := \{\sum_{i=1}^{s} g_i M_H(h_i) \mid g_i \in \Pi_{d-deg(h_i)}^{(H)}\}.$$

Analogously for an ideal $I \subset \Pi$,

$$V_d(I) := \{M_H(p) \mid p \in I,\ deg(p) = d\} \cup \{0\}.$$

We introduce an inner product $\langle \ . \ , \ . \ \rangle$ in Π, for instance, by the inner product of the (weighted) coefficient vectors, or by a strictly positive linear functional J and $\langle f, g \rangle := J(f \cdot g)$ if $\mathbb{K} \subseteq \mathbb{R}$ or $:= J(f\bar{g})$ if $\mathbb{K} = \mathbb{C}$. Then we can define orthogonal complements $W_d(h_1, \ldots, h_s)$ and $W_d(I)$ in $\Pi_d^{(H)}$. Hence

$$V_d(h_1, \ldots, h_s) \oplus W_d(h_1, \ldots, h_s) = \Pi_d^{(H)} \text{ and } V_d(I) \oplus W_d(I) = \Pi_d^{(H)}.$$

Let us consider a polynomial f of degree d. Then

$$M_H(f) \in V_d(h_1, \ldots, h_s) \oplus W_d(h_1, \ldots, h_s).$$

Let w_d denote its natural projection on $W_d(h_1, \ldots, h_s)$. This homogeneous polynomial can be computed by solving a finite linear system of equations because $\Pi_d^{(H)}$ has a finite dimension. Hence there are homogeneous polynomials $g_1, \ldots, g_s$ such that

$$f = w_d + \sum_{i=1}^{s} g_i h_i + f_1, \quad g_i \in \Pi_{d-deg(h_i)}^{(H)},\ f_1 \in \Pi_{d-1}.$$

We say f reduces to $w_d + f_1$ modulo $\{h_1, \ldots, h_s\}$ and call f_1 then the remainder of f.

In the reduction modulo $\{h_1,\dots,h_s\}$ the degree of the remainder f_1 is less than $deg(f)$. Hence this reduction can be applied recursively reducing f_{i-1} constructively to $w_{d+1-i}+f_i$ modulo $\{h_1,\dots,h_s\}$ for $i=1,\dots,d+1$ starting with $f=f_0\in\Pi_d$ and terminating with $f_{d+1}=0$, because the constant f_d is either in $V_0(h_1,\dots,h_s)$ or in $W_0(h_1,\dots,h_s)$. Combining these reductions modulo $\{h_1,\dots,h_s\}$, one obtains for f

$$f=\sum_{i=0}^{d}w_i+\sum_{i=0}^{d}\sum_{j=1}^{s}g_{ij}h_j,\quad g_{ij}\in\Pi^{(H)}_{i-deg(h_j)}.$$

Then $\sum_{i=0}^{d}w_i$ is uniquely determined by f, by $\{h_1,\dots,h_s\}$, and by the underlying inner product.

Definition 4. *Let $h_1,\dots,h_s\in\Pi$. We say $f\in\Pi_d$* reduces fully modulo *$\{h_1,\dots,h_s\}$ to $\sum_{i=0}^{d}w_i$ if every $w_i\in W_i(h_1,\dots,h_s)$ is constructed as described above. $\sum_{j=0}^{d}w_j$ is called the* normal form of *f* modulo *$\{h_1,\dots,h_s\}$, for short*

$$\mathtt{NF}(f,\{h_1,\dots,h_s\}):=\sum_{j=0}^{d}w_j.$$

If $\{h_1,\dots,h_s\}$ is not an H–basis of $I:=\langle h_1,\dots,h_s\rangle$, then $M_H(f)$ is not necessarily contained in $V_{deg(f)}(h_1,\dots,h_s)$, although $f\in I$. This means, that eventually the first homogeneous polynomial w_d is not 0 if $f\in I$. Hence at most if $\{h_1,\dots,h_s\}$ is an H–basis, then $\mathtt{NF}(f,\{h_1,\dots,h_s\})=0$. In fact, as quoted in [7] but shown already in [9], $\{h_1,\dots,h_s\}$ is an H–basis if and only if $\mathtt{NF}(f,\{h_1,\dots,h_s\})=0$ for every $f\in\langle h_1,\dots,h_s\rangle$.

Another characterization of H–bases given in [7] is as follows.

Theorem 3. *Let I be an ideal and $h_1,\dots,h_s\in I$. $\{h_1,\dots,h_s\}$ is an H–basis of I if, and only if, for all $d\in\mathbb{N}$,*

$$V_d(I)=V_d(h_1,\dots,h_s).$$

§4. On the Construction of H–bases

Macaulay proposed in [4] a procedure for computing H–bases of ideals given by a basis. However, his description was only by an example. He claimed "*This procedure is a general one*". But he needs in his example the computation of certain modules of syzygies. These can be constructed only in special cases or by computing first a G–basis and then applying techniques as in [1].

On the other hand, if an admissible term ordering $\prec$ is compatible with degrees,

$$deg(x_1^{\gamma_1}\cdots x_n^{\gamma_n})<deg(x_1^{\beta_1}\cdots x_n^{\beta_n})\ \Rightarrow\ x_1^{\gamma_1}\cdots x_n^{\gamma_n}\prec x_1^{\beta_1}\cdots x_n^{\beta_n},$$

then a G–basis with respect to $\prec$ is also an H–basis. Hence Buchberger's algorithm for computing G–bases also serves for computing H–bases. This seems a more direct access than via syzygies. However, if one wants to use H–bases instead of G–bases, this way is still a detour.

In case the number n of variables coincides with the number of given polynomials, then there is an easy test for H–bases as proved in [7].

Theorem 4. *Let $h_1, \ldots, h_n$ be n polynomials such that their maximal forms $M_H(h_1), \ldots, M_H(h_n)$ have only the point $(0, \ldots, 0)$ as common zero. Then $\{h_1, \ldots, h_n\}$ is an H–basis.*

For an arbitrary zero-dimensional ideal, i.e. for an ideal I such that the polynomials in I (equivalently: the polynomials in an arbitrary basis of I) have only a finite number of common zeros in $\overline{\mathbb{K}}^n$, $\overline{\mathbb{K}}$ the algebraic closure of $\mathbb{K}$, see [3], we present here a procedure which computes an H–basis from a given basis.

Procedure for computing H–bases.
In : $\mathcal{H}_o$, a finite polynomial set generating a zero-dimensional ideal I.
Out : $\mathcal{H}$, an H-basis of I.
Start: $\mathcal{H} := \mathcal{H}_o$, $d = 0$.
Loop: Check the finite dimensional linear vector space

$$V_d(\mathcal{H}) := \{\sum_{h \in \mathcal{H}} g_h M_H(h) \mid g_h M_H(h) \in \Pi_d^{(H)}\}$$

for linear dependencies. If $\sum_{h \in \mathcal{H}} g_h M_H(h) = 0$, then compute $p := \mathtt{NF}(\sum_{h \in \mathcal{H}} g_h h, \mathcal{H}))$. If $p \neq 0$, then enlarge $\mathcal{H}$ by p, and modify consequently $V_0(\mathcal{H}), \ldots, V_{d-1}(\mathcal{H})$. Lower d to the least k where $V_k(\mathcal{H})$ is changed and go to Loop. If for no linear dependency such p is nonzero, then then enlarge d by 1. If now $V_d(\mathcal{H}) = \Pi_d^{(H)}$ holds true, then return $\mathcal{H}$ otherwise go to Loop.

This informal description can be extended easily to a correct algorithm. One has to observe that the checking of $V_d(\mathcal{H})$ for linear dependencies needs a basis of the nullspace

$$\{(g_1, \ldots, g_s) \in \Pi_{d-deg(h_1)}^{(H)} \times \ldots \times \Pi_{d-deg(h_s)}^{(H)} \mid \sum_{i=1}^{s} g_i M_H(h_i) = 0\},$$

where $\mathcal{H} = \{h_1, \ldots, h_s\}$. If for every basis element $(g_1, \ldots, g_s)$ the normal form of $\sum_{i=1}^{s} g_i h_i$ is 0, then it holds for every element of the nullspace, i.e. for every dependency. As a byproduct of the basis computation one obtains $dim V_d(\mathcal{H})$. Then the test $V_d(\mathcal{H}) = \Pi_d^{(H)}$ reduces to a comparison of the dimensions because of $V_d(\mathcal{H}) \subseteq \Pi_d^{(H)}$.

For proving correctness and termination, we consider first $f := \sum_{h \in \mathcal{H}} g_h h$ with $g_h M_H(h) \in \Pi_d^{(H)}$ for all $h \in \mathcal{H}$. If $\mathtt{NF}(f, \mathcal{H}) = 0$, then especially $M_H(f) \in V_k(\mathcal{H})$ for a $k \leq d$, and hence

$$M_H(f) \in \langle M_H(h_1), \ldots, M_H(h_s)\rangle, \text{ where } \mathcal{H} = \{h_1, \ldots, h_s\}.$$

In case $p := \mathtt{NF}(f, \mathcal{H}) \neq 0$ either $M_H(f) \neq M_H(p)$ holds, i.e. again

$$M_H(f) \in \langle M_H(h_1), \ldots, M_H(h_s)\rangle,$$

or $M_H(f) = M_H(p)$ holds, i.e.

$$M_H(f) \in \langle M_H(h_1), \ldots, M_H(h_s), M_H(p) \rangle.$$

Therefore, if in the procedure d is increased by 1 (and $\mathcal{H}$ is updated), then for every $0 \neq \sum_{h \in \mathcal{H}} g_h h \in I$, $deg(g_h) + deg(h) \leq d$ the relation

$$M_H(\sum_{h \in \mathcal{H}} g_h h) \in \langle M_H(h_1), \ldots, M_H(h_s) \rangle$$

holds where again $\mathcal{H} =: \{h_1, \ldots, h_s\}$. This is our inductive hypothesis.

The ideal I has an H–basis, say $\{\varphi_1, \ldots, \varphi_m\}$. $\mathcal{H}$ is a basis of I. Hence every φ_i has a representation $\varphi_i = \sum_{j=1}^{s} g_{ij} h_j$, $g_{ij} \in \Pi$. If the inductive hypothesis holds for d, then one obtains at least for $d \geq M := \max_{ij} deg(g_{ij} h_j)$ that

$$M_H(\varphi_i) \in \langle M_H(h_1), \ldots, M_H(h_s) \rangle, \; i = 1, \ldots, m.$$

Hence for those d $V_d(\varphi_1, \ldots, \varphi_m) \subseteq V_d(\mathcal{H})$. But $V_d(\varphi_1, \ldots, \varphi_m) = V_d(I)$, since $\{\varphi_1, \ldots, \varphi_m\}$ is an H–basis of I. Therefore $V_d(\mathcal{H}) = V_d(I)$ for $d \geq M$. By the inductive hypothesis, also $V_k(\mathcal{H}) = V_k(I)$ holds for $k < M$. Hence $\mathcal{H}$ is an H–basis of I if we arrived at a $d \geq M$ in the procedure.

The ideal I has dimension 0. Then there is a D such that $V_d(I) = \Pi_d^{(H)}$ for all $d \geq D$, see for instance [3, Ch 9.4,Prop.6] and [3, Ch 5.3,Thm.6]. Hence for $d \geq max\{D, M\}$ one has $V_d(\mathcal{H}) = \Pi_d^{(H)}$. Thus in the course of the procedure, one arrives once, not knowing M, at a d_0 with $V_{d_0}(\mathcal{H}) = \Pi_{d_0}^{(H)}$. Then also $V_k(\mathcal{H}) = \Pi_k^{(H)}$ for all $k > d_0$. Therefore, for every polynomial $f \in I$ with

$$f = \sum_{i=1}^{s} g_i h_i, \; g_i \in \Pi_{k-deg(h_i)}^{(H)}$$

the assumption $M_H(f) \notin \langle M_H(h_1), \ldots, M_H(h_s) \rangle$ leads to $deg(M_H(f)) < d_0$. But then the inductive hypothesis gives a contradiction. Therefore the procedure gives no new $p \neq 0$ enlarging the set $\mathcal{H}$. This ensures termination.

An implementation of an algorithm based on this procedure and a complexity analysis is still a work under progress.

Acknowledgments. The second author was supported by the Deutsche Forschungsgemeinschaft with a Heisenberg fellowship, Grant Sa–627/6.

References

1. Adams, W. W. and P. Loustaunau, *An Introduction to Gröbner Bases*, Graduate Studies in Mathematics 3, AMS 1994.
2. Buchberger, B., Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal, (Doctoral thesis), Univ. Innsbruck, 1965.

3. Cox, D., J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, Springer Verlag, New York, 1992.
4. Macaulay, F. S., *The Algebraic Theory of Modular Systems*, Cambridge Tracts in Math. and Math. Phys. 19, Cambridge Univ. Press, 1916.
5. Möller, H. M., Gröbner bases and Numerical Analysis, in: *Groebner Bases and Applications, (Proc. of Conf. 33 Years of Groebner Bases)*, B. Buchberger and F. Winkler (eds.), Cambridge University Press 1998, 159–178.
6. Möller, H. M., and F. Mora: New constructive methods in classical ideal theory, J. of Algebra 100 (1986), 138–178.
7. Möller, H. M. and T. Sauer, H–bases for polynomial interpolation and system solving, Advances Computat. Math., to appear.
8. Möller, H. M. and T. Sauer, H–Bases II: Applications to numerical problems, *Curve and Surface Fitting: Saint-Malo 1999*, Albert Cohen, Christophe Rabut, and Larry L. Schumaker (eds.), Vanderbilt University Press, Nashville, 2000, 333–342.
9. Sauer, T., Gröbner bases, H-bases and interpolation, Proc. Amer. Math. Soc., to appear.

H. Michael Möller
Fachbereich Mathematik
der Universität
44221 Dortmund, Germany
hmm@mathematik.uni-dortmund.de

Thomas Sauer
Mathematisches Institut
der Universität Erlangen–Nürnberg
Bismarckstr. $1\frac{1}{2}$
91054 Erlangen, Germany
sauer@mi.uni-erlangen.de